# Privacy Regulations Challenges on Data-centric and IoT Systems: A Case Study for Smart Vehicles

Lelio Campanile, Mauro Iacono, Fiammetta Marulli* and Michele Mastroianni
*Dipartimento di Matematica e Fisica, Università degli Studi della Campania "L.Vanvitelli", Viale A. Lincoln, 5, Caserta, Italy*

Keywords: Internet of Things, Data Centric System, Internet of Vehicles, GDPR, Data Privacy, Data Security, Cyber-physical Systems, Blockchain.

Abstract: Internet of Things (IoTs) services and data-centric systems allow smart and efficient information exchanging. Anyway, even if existing IoTs and cyber security architectures are enforcing, they are still vulnerable to security issues, as unauthorized access, data breaches, intrusions. They can't provide yet sufficiently robust and secure solutions to be applied in a straightforward way, both for ensuring privacy preservation and trustworthiness of transmitted data, evenly preventing from its fraudulent and unauthorized usage. Such data potentially include critical information about persons' privacy (locations, visited places, behaviors, goods, anagraphic data and health conditions). So, novel approaches for IoTs and data-centric security are needed. In this work, we address IoTs systems security problem focusing on the privacy preserving issue. Indeed, after the European Union introduced the General Data Protection Regulation (GDPR), privacy data protection is a mandatory requirement for systems producing and managing sensible users' data. Starting from a case study for the Internet of Vehicles (IoVs), we performed a pilot study and DPIA assessment to analyze possible mitigation strategies for improving the compliance of IoTs based systems to GDPR requirements. Our preliminary results evidenced that the introduction of blockchains in IoTs systems architectures can improve significantly the compliance to privacy regulations.



## 1 INTRODUCTION

The Internet of Things (IoTs) has become a naturally included part of people's daily life; in the last decade, it introduced an effectively revolutionary system of technologies able to modify people's life, in sectors ranging from communication and services to transport, to health, from entertainment (Marulli et al., 2016),(Marulli and Vallifuoco, 2017) to our interactions with public institutions and government.

Anyway, as any relevant opportunity, also IoTs introduces a number of significant challenges, most of which are strictly related to users' security and privacy rights. Data are easily produced, exchanged and consumed by a multitude of services and devices but users don't know exactly where their data are physically stored and who can effectively access.

Indeed, both trustworthiness and security of data management systems have become issues of primary relevance; privacy and private data rights management has become so crucial to the point that governments were obliged to introduce specific regulations to ensure privacy preservation in any kind of transaction involving sensible information.

In May 2018, the European Union (EU) permanently introduced the General Data Protection Regulation (GDPR) (The European Union, 2016), a mandatory regulation on privacy data management applied across all the Members of EU. It has been applied directly to processing activities of personal data having relationships with any markets or territory in the European Union; from the date of its official entry to force, any breach of its provisions has been punishable by heavy sanctions, imposed by all Data Protection Authorities (DPAs) (Albrecht, 2016).

The GDPR was the latest step in the ongoing global recognition of the value and importance of personal information and it arised from the necessity to have a clear and unified regulation and protection measures in front of a fragmented digital market and the lack of enforcement in the field of data protection provisions. It represents a unified and directly applicable data protection law for the European Union

*Corresponding Author



Campanile, L., Iacono, M., Marulli, F. and Mastroianni, M.
Privacy Regulations Challenges on Data-centric and IoT Systems: A Case Study for Smart Vehicles.
DOI: 10.5220/0009839305070518
In *Proceedings of the 5th International Conference on Internet of Things, Big Data and Security (IoTBDS 2020)*, pages 507-518
ISBN: 978-989-758-426-8

which has replaced almost all of the existing Member States' provisions and which have been applied by businesses, individuals, courts and authorities without transposition into national law (Team, 2017).

The GDPR has conditioned not only the European data protection laws but the whole world, thus evidencing that the GDPR is currently one of the most significant data security law in the world.

So far, if the importance of respecting the GDPR regulations has been well understood by organizations,several are the concerns on its effective and correct implementation. It suggests a set of reference guidelines but it does not provide precise explanations or effective practices for transactions in order to be privacy preserving compliant.

For massive digital systems and services, as the IoTs based ones, the GDPR requirements introduced further criticalities, ranging from the design and implementation of novel robust and effective solutions for secure data management to the proof of compliance to mandatory regulations (Di Martino et al., 2019). So, data-centric and IoTs systems have been strongly impacted by these regulations and fully appropriate solutions haven't been still designed.

In this work, the impacts of GDPR privacy regulations on data-centric and IoTs systems design and applications are investigated. In particular, we focused on analyzing possible strategies to assess the level of compliance to specific privacy regulation, as the GDPR, that are exhibited by service infrastructures and applications, based on IoTs technologies.

To this aim, we considered a case study from the Internet of Vehicles (IoVs), describing the scenario of an accident occurring among smart vehicles. Such case study was functional to perform a pilot study for analyzing mitigation strategies to better meet trustworthiness and GDPR privacy constraints. When an accident occurs, privacy concerning data are typically exchanged among several stakeholders, ranging from the people involved in the accident (e.g., the vehicles drivers or witnesses) to the insurance companies and/or the law enforcements.

In such a scenario, data are quickly exchanged by multiple stakeholders and no one can exactly ensure who will be enabled to access to that data, referring people's locations, involved people's identities and personal data. So, data can't be exchanged by common communication services and means but need to be managed in appropriate ways, able to ensure the freedom and the right of anonymity besides the effective stakeholders.

To the aim of assessing any possible mitigation strategy, we discussed two variant scenarios for the proposed case study, considering, in turn the absence and the presence in introducing the use of a blockchain in the data exchanging and managing loop of an IoVs application.

Indeed, we performed a comparison between the two scenarios (with and without adopting the use of the blockchain), evidencing the benefits obtained by using blockchains, in terms of an increased compliance to the privacy preserving goal prescribed by GDPR. The degree of compliance was defined as the estimation of the risk level occurring for an IoTs application to be in privacy regulation violation region, with or without introducing the blockchain in the data managing loop.

Finally, the pilot study provided in this work is functional to further studies for defining a systematic metrics framework for assessing privacy regulations compliance for practical IoTs service applications.

The rest of the paper is organized as follows: Section 2 presents related works, section 3 describes the considered case study for an internet of vehicles application; in section 4 a blockchain based and GDPR-compliant solution we proposed is described; in section 5 the results obtained by the proposed approach are discussed and section 6 concludes the work by descring the insights of our study and the further directions for improving this investigation.

## 2 RELATED WORKS

Internet of Things (Atzori et al., 2010) primarily means a world-spanning information fabric built on a wide range network computing environment made of smart devices, able to provide smart solutions, using internet for communicating automatically among each other without requiring any human intervention. Such intelligent devices perform actions basing on the information they get from other connected devices anytime and anywhere; this is how they perform their designated tasks intelligently by deciding in real time (Solangi et al., 2018). As the next generation of the Wireless Sensor Networks, IoTs systems and infrastructures behaviors and performances can be also simulated by advanced network simulators as the NS-3 (Campanile et al., 2020), provide with tailored libraries for supporting the various IoTs scenarios.

However, beyond several fantastic opportunities, IoTs also presents a number of significant challenges. The growth in the number of devices and the speed of that growth presents challenges to our security and freedoms that require to develop policies, standards, and governance able to shape this development without stifling innovation. Various applications of IoTs, like smart home and buildings, smart health-

care, smart vehicles and smart appliances, e.g., have anticipated yet security issues due to authentication and data integrity and, in the recent past, an increasing number of researches have focused on security, data privacy and trust concerns related to IoTs infrastructures. One of the main factor causing security and privacy pitfalls and faults in the IoTs ecosystem, as discussed in (Maple, 2017), can be found in the lack of standardisation and regulation, since the IoT has evolved using a wide range of core technologies from a number of key visions, through developments by distinct, often disparate communities. Further, these developments have been made in different application areas, often using specific and proprietary standards. This diffuse nature of development has led to an inevitable lack of harmonisation and shared vision, hampering standardisation and effective regulation, which has left technicians and users without the necessary information and control to, service, update, and address problems created with devices and services. The lack of coherence, oversight, understanding, and protocols means that security risk analysis, risk assessment, and countermeasure implementation (Cimino et al., 2020) are much more difficult tasks than they would be with a more directed and coordinated development path. (Maple, 2017) provides a deep discussion on the security and privacy challenges in the IoT, illustrated through a number of key applications. In this discussion, the securing principles of systems are also discussed, from the CIA of information security (confidentiality, integrity, and availability), to the five keys of information assurance (confidentiality, integrity, availability, authenticity, and non- repudiation) (Frattolillo et al., 2009) and the Parkerian Hexad (confidentiality, integrity, availability, authenticity, possession, and utility) (Parker, 1983). Research works discussing security considerations relating to complex cyber-physical (as opposed to information) and IoT systems vary in which principles they adopt. The majority of researchers restrict consideration to the CIA. The Parkerian Hexad usefulness remains the subject of debate among security professionals (Sattarova Feruza and Kim, 2007). Others studies also include robustness, reliability, safety, resilience, performability, and survivability (Sterbenz et al., 2010). challenges.

As for the authentication within the IoT, it is a critical point too, since without appropriate authentication the confidentiality, integrity, and availability of systems could compromised. This is the reason for an adversary to be able to authenticate as a legitimate user, to access to any data that the user has, and be able to see (compromising confidentiality), modify (compromising integrity), and delete or restrict availability (compromising availability) in the same way that the user can. The authentication and identification of users in the IoT remains a significant challenge. Currently, username/password pairs are the most common form of authentication and identification of users in electronic systems. However, the vision of the IoT as ubiquitous will eliminate many of the physical interaction interfaces through which usernames and passwords are passed.

Furthermore mobility, privacy, and anonymity require further analysis and research because those IoTs systems featuring mobile services will have users passing through different architectures and infrastructures owned by different providers. Managing the identity of users in mobile, multiply owned and heterogeneous environments is strongly challenging.

As for the security and privacy challenges in the IoTs environments, in (Mena et al., 2018) is provided a survey from the perspective of the adopted technologies and architecture. This research is focused on IoT intrinsic vulnerabilities and their implications to the fundamental information security challenges in confidentiality, integrity, and availability. From a wide range exploration of the IoT architecture security related issues, concluding that all layers of IoT architectures are involved in security issues: data access and authentication, phishing attacks, malware attacks (Mercaldo and Santone, ),(Casolare et al., 2019), malicious scripts are all examples of threats for the IoT application layer are; in the IoT network layer, network congestions, devices interference, eavesdropping attack, routing attack, denial of service,attack node jumping in WSN and heterogeneity problem are examples of security issues.

## 2.1 The GDPR

As for the privacy issues (YANG et al., 2019), the scenario was furtherly complicated by the introduction of mandatory regulations for data protection in all the world countries. As for the European Union, it introduced the General Data Privacy Regulation (GDPR) (Voigt and Von dem Bussche, 2017) that currently requires all data controllers and processors that handle the personal information of EU residents to adopt and implement appropriate technical and organisational measures to ensure the ongoing confidentiality, integrity, availability and resilence of processing systems and services or face heavy financial penalties. The GDPR is the latest step in the ongoing global recognition of the value and importance of personal information. Such a recognition has obtained only recently the appropriate attention, because of the increasing number of cyber thefts and the misuse of per-

sonal data by governments and corporations, that expose EU citizens to significant personal risks. Thus, the need of new law clarifying the data rights of EU citizens to ensure the appropriate level of EU-wide protection for personal data, led to the composition of the GDPR, that applies across all the Member Countries of the EU but involves any organisation anywhere in the world providing services into the EU that involve processing and managing personal data (Team, 2017). Consequently, technological smart services and infrastructures (e.g., smart cities) dealing with sensitive data have been strongly impacted since they must ensure individual privacy and security in order to ensure that people will going on participating. Sensitive data (*special categories of personal data*, as defined in Article 9 of GDPR), which are generated by a plurality of subjects or connected to the interest of a plurality of subjects, have to be collected, maintained and kept unaltered but inaccessible until legally authorized. Internet-of-Things (IoT) has become a naturally included part of our daily life with billions of IoT devices collecting data through wireless technology and able to interoperate within the existing Internet infrastructure. Moreover, the new fog computing paradigm allows storing and processing data at the network edge or anywhere along the cloud-to-endpoint continuum, thus overcoming the limitations of IoT devices. IoT, fog computing and smart cities are all representative of Data Centric Systems and paradigms, needing of being provided by people's partecipation and data in order to be working. If people are reluctant to participate, the core advantages of all data centric infrastructures will dissolve. Security and privacy issues, such as secure communication, authentication and authorization, information confidentiality and integrity, introduce destabilizing and costly disruptions for data centric systems.Moreover, the widely distribution of millions of smart devices, located in different areas, increases the risk of being compromised by some malicious parties. Actual challenges that Data Centric Systems are obliged to issue and manage includes privacy preservation with high dimensional data, securing networks with large attack surface, establishing trustworthy data sharing practices, properly utilizing artificial intelligence, and mitigating failures cascading through the smart network, in order to avoid people deceiving and data protection regulations penalties.

## 2.2 Blockchain Applications to IoTs Systems

Blockchain, or distributed ledger technology, is an emerging platform that is designed to support transactions services within a multi-party business network, hopefully ensuring for all involved partners significant cost and risk reductions. In the last few years, this technology has started to be used in a wider range of application, including Internet of Things (IoT). A blockchain enables IoT devices to send data for inclusion in a shared transaction repository with tamper-resistant records, and enables business parties to access and supply IoT data without the need for central control and management Many researchers have addressed the problem of integrating blockchain with IoT.

Blockchain (BC) technologies are well known for their feature of data immutability (Zheng et al., 2019). A BC, in fact, consists of an open and distributed ledger running over a peer-to-peer (P2P) network that can manage transactions for multiple entities efficiently and in a verifiable and traceable way, thus without recurring to the intervention of a middleman.

A blockchain is made of a list of blocks recording a transaction and linked using cryptography. Each block contains a timestamp represented by a cryptographic hash of the previous block, where the data of the transaction are represented as a Merkle tree. The tamperproof nature of blockchain comes from the fact that, once consensus is reached and a block is committed, the data in the block cannot be altered retroactively without alteration of all subsequent blocks, which requires the consensus of the majority.

Furthermore, BC can follow both a centralized architecture and a decentralized one but the fundamental problem of scalability keeps the blockchains operating in beta and alpha mode in business and industry.

All these security and performance concerns are due to the "Scalability Trilemma", where a blockchain system tries to offer scalability, decentralization and security, without compromising any of them. Decentralization is the core property that enables the censorship-resistance and permissionless features. Scalability is the ability to process transactions on a network with increased size.

Security is an important component that guarantees the immutability of the ledger and its resistance to general cyber attacks.

As for the Internet of Things (IoTs) applications (Wang et al., 2019), BC technologies have the potential to address the data security concern in IoT networks. While providing data security to the IoT, Blockchain also encounters a number of critical challenges inherent in the IoT, such as a huge number of IoT devices, non-homogeneous network structure, limited computing power, low communication bandwidth, and error-prone radio links. Traditional solutions are based on registries owned and managed by

trusted third-party or Authorities representing public regulators with generally complex, anyway certified insertion or updating procedures. The regulator can actually manage this registry as a whole, if the frequency of operations can be actually handled, as a system of partial registries, articulating the distribution of operations on the basis of proper criteria (e.g., according to a well defined geographic partitioning), or acting as a top authority of a tree-organized delegated registries that are provided by third parties according to specific agreements and duties, that allows online legal identification of persons by using tokens issued by third parties that can verify credentials. However, these solutions represent a logical single point of failure, when non distributed, a cost (including technical, organizational and legal components), when maintained by public regulators, or a strong delegation choice, when involving commercial third parties that may be also providing other non-regulated services, potentially exposed to security breaches. Furthermore, the spreading of the Internet of Vehicles, that enables a large number of services ranging from vehicle-to-vehicle communication to autonomous drive and platooning, to the interaction with smart roads (Karpiriski et al., 2006), makes accessible the use of a blockchain for traffic monitoring, security and safety support and the solution of legal issues because of the native integration of sensing and computing on board and the possibility of capturing and cross compare the local state of the traffic. The fact that different stakeholders have a legitimate interest in ensuring the availability of all relevant information to support their internal processes, to discharge responsibilities, to certify operations and to disambiguate critical situations with a lower level of risk and reducing the need for unilateral preventive solutions makes sharing such a blockchain infrastructure viable and convenient. In the more specific field of integrating Blockchain technology and Internet of Vehicles (IoV) (Odiete et al., 2018), propose a fully decentralized architecture that combines blockchain and traditional distributed database to gain additional features such as efficient query and retrieval of metadata stored on the blockchain. (Truong et al., 2018) assert that before making any transactions it is crucial to evaluate trust between participants for reducing the risk of dealing with malicious peers, and is proposed a trust-based IoV model including a system architecture, components and features. (Arora and Yadav, 2018) present an authentication and secure data transfer algorithm, in the IoV framework using blockchain technology, to ensure secure information communication and to improve the efficiency. Furthermore, Regarding blockchain and IoT integration, (Reyna et al., 2018) analyze the challenges emerging from the integration of IoT and blockchain, presenting possible ways of integration and platforms that are integrating IoT and blockchain in a general context. (Christidis and Devetsikiotis, 2016) describe how a blockchain-IoT combination may facilitates the sharing of services and resources leading to the creation of a marketplace of services between devices and allows to automate in a cryptographically verifiable manner several existing, time-consuming work ows.

# 3 THE PROPOSED PILOT STUDY FOR PRIVACY COMPLIANCE

The rise of the Internet of Vehicles, that enables a large number of services ranging from vehicle-to-vehicle communication to autonomous drive and platooning, to the interaction with smart roads, makes accessible the use of a blockchain for traffic monitoring, security and safety support and the solution of legal issues because of the native integration of sensing and computing on board and the possibility of capturing and cross compare the local state of the traffic. Moreover, the integration of production and maintenance information may be soon seamlessly available with the same technology, as literature reports the appearance of blockchains to support supply chain in manufacturing sectors like aerospace and automotive. The fact that different stakeholders (insurances, manufacturers, garages, black box service providers, smart road concessionaires) have a legitimate interest in ensuring the availability of all relevant information to support their internal processes, to discharge responsibilities, to certify operations and to disambiguate critical situations with a lower level of risk and reducing the need for unilateral preventive solutions makes sharing such a blockchain infrastructure viable and convenient. Moreover, a careful design of the data management and signature mechanism may also provide the possibility of implementing different levels of privacy, so to have non-personal but commercially significant data certified and accessible by the authorized party to produce additional information-driven services in the interest of the final users.

## 3.1 The Proposed Case Study for the Internet of Vehicles

The case study we considered for the aim to investigate mitigation strategies for major compliance of IoTs based systems to privacy regulations, was taken from the Internet of Vehicles (IoVs) application do-

main and the specific privacy regulation we considered is the European Union GDPR. The main idea behind our pilot study consists is the fact that several are the stakeholders interested in both managing the information about the behavior of the vehicles and defending themselves from possible privacy rights violations deriving from data detention. Indeed, the uncontrolled tracking of data concerning events related to personal or commercial information can't be allowed,while the availability of a non-contestable registration and partial access to anonymized information with sharing of responsibility is needed. In the proposed case study, we adopted asymmetric key pairs both to sign information and assure the requirement of non repudiation, and to encrypt sensitive information, in order to allow the authorized access only to the legitimate involved stakeholders or to disclose information to the authorities when a law judgement procedure occurs. Furthermore, we adopted the usage of blockchains to share responsibility, guarantee and verify immutability of information, beyond a lower level of risk and protection costs, without disclosing sensible information.

## 3.2 Internet of Vehicles Scenario

The set of stakeholders we considered in our study is depicted in Figure 1and briefly characterized as follows:

- vehicles drivers;
- vehicles manufacturers;
- insurance companies;
- smart road dealerships;
- car parking garages;
- public registry authorities;

Drivers are the effective users of vehicles. Users represent citizens that own or drive private vehi- cles, while company users represent people that drive vehicles belonging to the farm they are working for; finally, vehicle rental customers represent people that drive vehicles belonging to a company that rented them for a short or long period to them. In the first case, a user is not expected to be a participant of a blockchain on his own, but he interacts with the blockchain by his vehicle, and the vehicle is expected to be equipped with a black box and a black box service provider acting as a proxy; the user may also interact with the blockchain for bureaucracy through another participating blockchain service provider, and the black box will directly provide services to a vehicle as well. In the second case, vehicles are under the responsibility of a fleet manager whose company participates the blockchain to track vehicle events that will be disclosed by the authority if needed, to enact defensive strategies and build a non repudiable log to prove drivers' responsibilities, and to collect anonymized data about own fleet for management, maintenance and further optimization reasons.

A vehicle manufacturer produces vehicles and performs maintenance in the warranty period. So, it can be considerate a legitimate interest allowing this stakeholder to collect diagnostic data about the circulating vehicles it has produced in order to collect defects affecting safety of drivers, circulation and other people or items. Anyway, data should anonimyzed in order to be not linkable to the owners identities. No other stakeholder should be able to obtain this information, while its correlation to the sense status of the car should be noncontestable, to let authorities understand as better as possible if any accident might occur. Such a stakeholder has the financial and technical capability needed to participate in a blockchain, that can be used to track all data related to its supply chain and production processes.

A insurance company has no right to access any information about the vehicle until no events covered by an insurance policy occur. This stakeholder has a legitimate interest in holding a copy of all encrypted data to be assure the veracity of information used in case of disputes or trials. In this perspective, it can take advantage in participating in the blockchain, even if data are anonimyzed, if the authority provides plain data and a policy to check them against the blockchain without violating the users' privacy.

A smart road dealership tracks data about vehicles passing by road it grants for, for limited time usage or based on statistical aggregations over long periods. It aims for improving safety, allowing autonomous drive, traffic management and implementing strategies for improving efficiency. It gets systematically in control of a large quantity of data about drivers' behavior, vehicle movement and traffic patterns, possibly with a chance for managing enough data to be potentially used to violate privacy and mine sensitive information: thus, it is interested both in a third-party certification of the veracity of information, eventually requested by authorities when incidents occur and in sharing responsibilities. Both the services can be privided by participating in the blockchain, without disclosing data and with the possibility of processing anonymized data with the aid of law authorities as a compensation for the production of public interest data.

A vehicle parking garage works on vehicles to perform scheduled or extraordinary maintenance, poten-

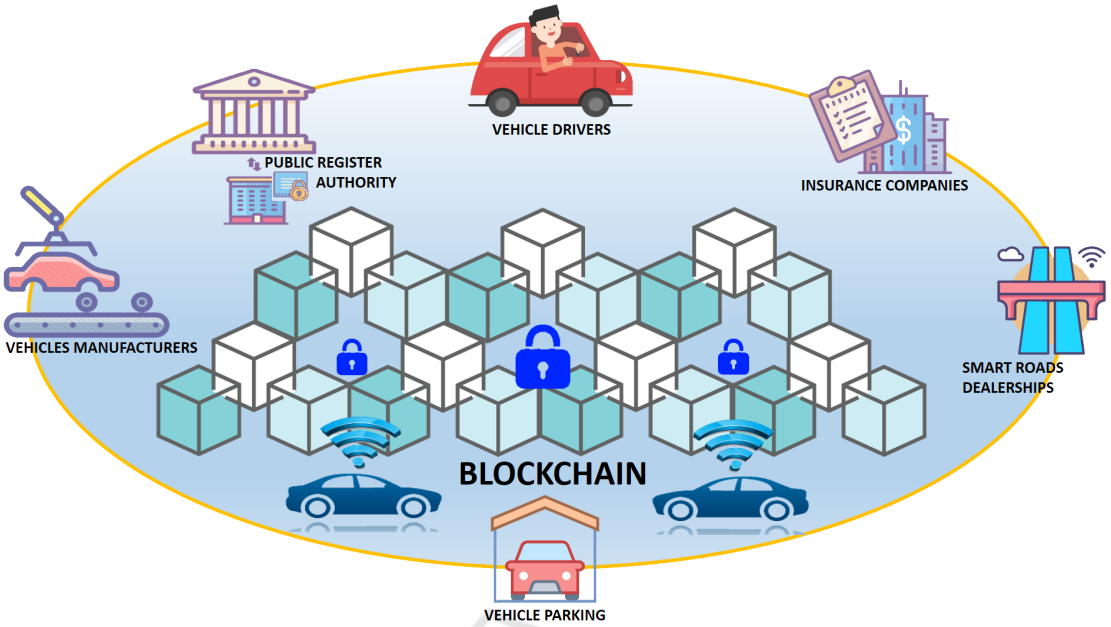


Figure 1: Stakeholders in the Blockchain Based IoVs.

tially impacting the safety of the vehicle. This stakeholder, in the person of the owners, has a legitimate interest in reliably tracking operations to limit own responsibility, while the manufacturer could benefit from being updated with maintenance events occurring outside the boundaries of its own branches. Anyway, participating the blockchain may be not technologically possible for small business organizations, like garages, so a trusted proxy could participate the blockchain and provide services to garages.

Finally, the public registry represents the authority managing accesses to sensitive data. It participates the blockchain on a peer role, evenly provided with the rights to access all data on behalf law authorities or when insurance companies signal the starting of a procedure involving their ordinary activities. Participating the blockchain allows the public registry to retain a copy of all data without the duty of directly managing all registrations from all parties and without the responsibility of being the only official repository of all data. The public registry is here considered a trusted party operating by applying all legal requirements, including systems security and enforcing all law constraints over the blockchain, that impacts over all participants. The public registry has the responsibility of keeping all public cryptographic keys used in conjunction with the blockchain, the private key needed to access to sensitive data logged by the blockchain and the mapping between identities of users and vehicles and between vehicles and blackboxes.

# 4 THE GDPR-COMPLIANT BLOCKCHAIN BASED SOLUTION

The GDPR is always applicable when personal data are managed and processed. The more significant rights for users include several rights, among which the ones to access own data or data about oneself, to change or fix wrong and missing information about oneself, to request the erasure of data about oneself ("right to be forgotten"). Some of these rights seem to be forged to shape the features of blockchain systems, as the data added to a digital ledger, without the possibility to change or remove it after its writing. Furthermore, blockchain-based solutions may ensure more security features when compared with other solutions: there are no single point of failure, since it can improve fault tolerance and makes hacking data very difficult. So, a blockchain-based solution may be profitably used in a GDPR-compliant environment and taking into account some remarkable constraints. In order to ensure this constraint, all participants (nodes) holding personal data has be known by all the users; all of them need to be defined as joint controllers (see Articles 24 and 26, GDPR). The most suitable model of blockchain fitting these constraints follows a private/permissioned paradigm (van Geelkerken and Konings, 2017). Moreover, in order to ensure full GDPR compliance, one of the main challenges for blockchain developers is to comply with

the right to rectification (Article 16) and the right to be forgotten (Article 17). Users have the right to request for rectification of their own personal data, and, in some conditions, may request the erasure of own personal data. Personal data must also be erased when the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed (Article 17), even if the user did not request for it. Accordingly to these constraints, the proposed blockchain based solution fits the following requirements:

- it is private and permissioned;
- data modifications are allowed but each update is recorded in an additive block;
- data erasing is allowed by encrypting blocks and deleting the key.

#### 4.0.1 Use Cases for the Compliance Analysis

The main use cases considered in the study for discussing the compliance analysis to GDPR are the listed and briefly described in the following:

- new vehicle registration: users can interact with the system for registering a new vehicle before its first usage, to activate all required procedures enabling the tracking by preserving his/her privacy, to link the vehicle with its black box and to bind his own identity to the vehicle and the black box.
- new maintenance registration: a vehicle manufacturer could register maintenance when the vehicle is maintained while the warranty period or at controlled garages,and should be able to access all diagnostic information about each single produced car.
- a smart road concessionaire monitors events occurring while vehicles ride its granted road and should be able to register them officially, including information that relates events with the involved vehicles.
- an insurance company interacts with its customers outside the system for the most of the cases, because all commercial activities, communications and service requests happen in presence, by phone or through the internet.
- The public registry needs to access data on authority request with full availability of contents, to register pseudo identities, being the only entity that keeps the correspondence between pseudo-identities and vehicles, handles insurance requests by providing data about involved parties about a limited time frame and verification means in case of incidents, encrypts obsolete data to delete them automatically when re- quested by GDPR or encrypts data about a user to delete them on user request as allowed and prescribed by GDPR.

## 5 THE GDPR COMPLIANCE ANALYSIS

Ensuring privacy preservation and data protection is mandatory for Government agencies and enterprise organizations that require personal data of their customers for providing IT services. Article 35 of the GDPR refers to new technologies and prescribes data processing by these novel technological could turn into a high risk to the rights and freedoms of natural persons. So, each controller shall carry out a Data Protection Impact Assessment (DPIA), that aims to conduct a systematic risk assessment in order to identify system vulnerabilities and privacy threats, thus prescribing and imposing technical and organizational controls to mitigate those threats. Article 35 prescribes that DPIA shall in particular be required in the case of:

- a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- processing on a large scale of special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offenses referred to in Article 10;
- a systematic monitoring of a publicly accessible area on a large scale. The case we are studying, a car tracking system, relates both to item a. and c. cited before; so, according to GDPR, DPIA must be carried out. In the following subsections, a DPIA is performed on our system compared with an hypothetical state-of-the-art cloud-based system to emphasize advantages of our design choice. In this paper we do not wish to draw up a complete impact assessment, as it would need more technical and organizational details. Our aim is to show which advantages may be obtained by using a blockchain- based framework instead of a simple cloud-based solution, and only some significant threats are evaluated.

### 5.1 DPIA and Risk Analysis

Risk is defined, according to standard ISO 31000:2009, as the effect of uncertainty on ob-

jectives, and it is often quantified as the probability of occurrence of an adverse event, times its negative impact (severity). Information Security Risk Assessment (ISRA) is commonly classified according to three criteria: Quantitative, Qualitative, and Hybrid (Semi-Quantitative) (Shameli-Sendi et al., 2016).

Quantitative risk assessment is based on objective measurements; its results are expressed in management-specific metrics (i.e. monetary value, percentages and probabilities). The main problem with quantitative assessment is the heavy effort required in terms of time-consuming and it is biased by the quality of information. Due to these limitations too, this approach is not so easy to be practically implemented. Alternatively, the qualitative approach is the most common into information security risk assessment processes, where many organizations find it adequate for their needs (Landoll and Landoll, 2005). In a qualitative approach, classes are introduced to show the impact and probability of a particular scenario. The qualitative approach is widely used, because of the lack of very accurate historical data for computing the impact and probability of occurrence of risks; it results to be also much easier to understand and implement, the calculations involved are simpler and the evaluation of assets,threats and vulnerabilities is simplified (Landoll and Landoll, 2005). Hybrid approach represents a combination of the two previous approaches (Shameli-Sendi et al., 2016).

## 5.2 Risk Assessment Metrics

In this study we adopt a qualitative approach based methodology, as the one adopted by CNIL, the French Data Protection Authority; this methodology is also recommended by many EU Data Protection Authorities, including the Italian Authority for Personal Data Protection. In addition to the methodology, CNIL also developed a software tool (PIA) to support managers in drawing up DPIA.

The CNIL-PIA software provides classifications accord to risk sources:

- internal human resources (employees, managers, e.g.);
- external human resources (business partners, visitors, maintenance staff);
- non-human sources (malicious code such as viruses, worms etc., earth-quake, malicious users);

and according to feared events:

- disappearance of personal data.
- illegitimate access to personal data;
- unwanted modification of personal data;

Severity and likelihood are the qualitative metrics estimated for performing the risk assessment. Severity represents the magnitude of a risk and it is primarily estimated in terms of the extent of potential impacts on data subjects, taking account of existing, planned or additional controls. Its value is estimated by considering the potential damage occurred to the user/data subject. The scale for estimating severity ranges among the following levels:

- 1. negligible: not affected/few inconveniences (loss of time, unsolicited mail);
- 2. limited: significant inconveniences (denial of access in services, additional costs);
- 3. significant: significant consequences (fraud, damages to property);
- 4. maximum: significant/irreversible consequences (inability to work, loss of evidence in the context of litigation, loss of access to vital infrastructure).

The severity level may be lowered by including additional factors to contrast identification of personal data, such as encryption, pseudoanonymization, anonymization, and so on.

Likelihood represents the feasibility of a risk to occur and it is primarily estimated as the level of vulnerabilities concerned to the supporting assets and the level of capabilities of the risk sources to exploit them, evenly considering existing or planned controls. The classification scale for likelihood is related to the feasibility of the occurrence of a risk, for each selected risk sources that is able to exploit the properties of supporting assets. As for the severity metric, also the likelihood can be estimated according four ranged values:

- 1. negligible: not (apparently) possible (e.g. theft of paper documents stored in a room protected by a badge reader and access code);
- 2. limited: difficult (e.g. theft of paper documents stored in a room protected by a badge reader);
- significant: possible (e.g. theft of paper documents stored in offices);
- 4. maximum: easy to breach (e.g. theft of paper documents stored in the public lobby).

The likelihood level may be decreased by including security enforcement components, as firewalls, logging and monitoring services and so on.

## 5.3 DPIA Results and Discussion

In order to perform an impact assessment analysis and compare the proposed blockchain solution, we con-

sidered a second system, consisting of a hypothetical IoT and cloud-based solution, accomplishing the following specifications:

- backup and business continuity features, more than two different physical sites;
- cryptography of all data;
- use of digital signature;
- operation security compliant with a well-known standard (eg. ISO 27001);
- archiving and tracking feature for all transactions.

We also considered some potential impacts burdening to users/data subjects. The choice of impacts is related to the specific kind of data involved. The impacts taken into account, and the related severity according to CNIL-PIA, are represented by:

- cost rise (severity = 2);
- targeted online advertising on a confidential aspect (severity = 2);
- fraud (severity = 3);
- loss of evidence in the context of litigation (severity = 4).

Threats selection was mainly based on the data retrieved on Verizon 2019 Data breach Investigation report (Ashraf, 2019) and EY Global Information security Survey 2018-19 (Feng and Wang, 2019). These threats can lead to illegitimate access (I), unwanted modification (U) of or disappearing (D) of Personal Data and are listed as follows:

- hacking: it is the most frequent threat, circa 54% (Verizon Enterprise (2019)) (I), (U);
- use of stolen credentials: almost 30% of threats (Verizon Enterprise (2019)) (I), (U);
- privilege abuse: circa 10% of threats (Verizon Enterprise (2019)) (I), (U), (D);
- DDoS attacks: almost 60% of attacks in incidents, usually this kind of attack does not lead to a data modification or illegitimate access, but leads to loss of availability (data disappearing), which is considered a breach event in GDPR (D);
- natural disasters: although they are infrequent events (EY appraises circa 2% of total breaches (EY (2019))), they are taken into account because they can lead to a severe data loss (D).

In case of the considered cloud based solution, the value of Risk Severity and Likelihood are estimated as follows:

- illegitimate access to personal data: the involved impacts are cost rise, confidential aspects and fraud, so severity is placed at 3 (significant), because this is the maximum level of severity in this set. Likelihood is placed to 3 (significant), due to the relatively consistent probability of occurrence of the event;
- unwanted modification of personal data: all impacts are involved, so severity is 4 (maximum). Likelihood is placed to 3 (significant), due to the relatively consistent probability of occurrence of the event;
- disappearance of personal data: the impact involved are cost rise and loss of evidence, so severity is 4 (maximum), but we think that state-of-the-art countermeasures such as business continuity may lower the severity to 3. Likelihood is placed to 2 (limited), due to the low probability of occurrence of the event.

The insights of our DPIA analysis evidenced in are shown in Fig. 2. In this picture we used the green region and the red region to represent compliance and non compliance regions. For the cloud based solution we observed that DPIA leads to a "High Risk" evaluation and, according to Article 36 of GDPR, it is mandatory for the data controller to consult the supervisory authority before processing. In the case of the assessment for the blockchain-based solution, the values were estimated as follows:

- illegitimate access to personal data: the involved impacts are cost rise, con
  dential aspects and fraud, so the severity is placed at 3 (significant). In this case, severity is lowered due to the large-scale use of pseudonymization. Users' data are all exchanged using pseudo identities; vehicle data are anonymized and decoupled from user identfiers and pseudo identity data. Likelihood is also lowered to 2 (limited): the risk due to stolen credential and privilege abuse is significantly lower, because the only credentials which may see/modify personal data are those assigned to the public registry and no sysadmin or black box manager has any access. Hacking probability is also decreased, due the extremely low probability of a success in enacting a hacking attack that overtakes 50% + 1 of the blockchain nodes;
- unwanted modification of personal data: all impacts are involved, so severity is 4 (maximum). Also in this case, the severity is mitigated due to the large-scale use of pseudonymization. Users' data are all exchanged using pseudo identities; vehicle data are anonymized and decoupled from user and pseudo identity data. In this case, the severity and the likelihood values are lowered to

2 (limited). The risk due to the fact that stolen credential and privilege abuse events likelihood is significantly lower, because the only credentials which may see/modify personal data are those assigned to the public registry and no sysadmin or black box manager has any access. Hacking probability is also lower, as for the previous point;

- disappearance of personal data: the involved impacts are cost rise and loss of evidence, so severity is 4 (maximum), but we think that state-of-the-art countermeasures such as business continuity may lower the severity to 3. Likelihood is also lowered to 1 (negligible): the event has a low probability to occur, and the large number of nodes involved make almost impossible to lead to a permanent data loss.

Summarizing up the DPIA results, in Figure 2 red dots represent the cloud-based solution while blue dots represent the blockchain-based solution. Is noticeable that in the case of the blockchain-based solution all the tree values (I, U and D) are placed in the ”green” area. This means that, according to Article 26, DPIA leads to ”medium risk” evaluation and there is no need to consult the supervisory authority before processing.


Figure 2: Risk Metrics Estimation for the Blockchain-based solution.

## 6 CONCLUSIONS

In this article we have discussed a pilot study aiming to perform a preliminary analysis on the impacts of data privacy regulations, on data-centric and IoTs service infrastructures and applications. Indeed, one of the major challenge for security and privacy assurance in the IoT, posed by the introduction of privacy regulations is still represented by the definition and the standardization of tangible specifications that could be transformed in software and hardware requirements. Arguably the most fundamental challenge, in the close future, is to encourage standardisation and coordination in the IoT systems design, by taking into account all conflicting views of the sevearl stakeholders on the IoT and the difficulties related in gaining consensus and trust between parties with different interests, goals and visions. In this perspective, we investigated any mitigation strategies for assessing the compliance of IoTs and data-centric systems to privacy regulations, as the European Union GDPR, putting also the basis for further studies aiming to identify a systematic metrics framework to guide the practical design and implementation of IoTs and data-centric systems and applications to be effectively compliant to data privacy regulations and for measuring the compliance level, in order to prevent violations. We performed our study by starting from a case study taken from the Internet of Vehicles and we aimed to observe if the introduction of a blockchain in the data managing loop could mitigate the impact of data privacy regulations on the system, in terms of compliance to its mandatory requirements. The degree of compliance was defined as the estimation of the risk level occurring for an IoTs application to be in privacy regulation violation region, with or without introducing the blockchain in the data managing loop and was performed by the means of a DPIA. As a preliminary result of this study, we evidenced that the introduction of a blockchain in the data loop could mitigate part of these problems and improve the position of systems and application from the boundaries of the region in which a regulations violation could occur. We explored this possibility in this work and we obtained, as we expected, promising insights that will be further investigate in the proceeding of the study we have proposed in this work.

## ACKNOWLEDGEMENTS

This work is part of the research activities realized within the project ”Attrazione e Mobilità dei Ricercatori” Italian PON Programme (PON_AIM 2018 num. AIM1878214-2).

## REFERENCES

Albrecht, J. P. (2016). How the gdpr will change the world. *Eur. Data Prot. L. Rev.*, 2:287.

Arora, A. and Yadav, S. K. (2018). Block chain based security mechanism for Internet of Vehicles (IoV). In

*Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT)*, 2018, pages 267–272.

Ashraf, C. (2019). 5g and mixed reality glasses help change how we see the world. verizon enterprise tech.

Atzori, L., Iera, A., and Morabito, G. (2010). The internet of things: A survey. *Computer networks*, 54(15):2787–2805.

Campanile, L., Gribaudo, M., Iacono, M., Marulli, F., and Mastroianni, M. (2020). Computer network simulation with ns-3: A systematic literature review. *Electronics*, 9(2):272.

Casolare, R., Martinelli, F., Mercaldo, F., and Santone, A. (2019). A model checking based proposal for mobile colluding attack detection. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 5998–6000. IEEE.

Cimino, M. G., De Francesco, N., Mercaldo, F., Santone, A., and Vaglini, G. (2020). Model checking for malicious family detection and phylogenetic analysis in mobile environment. *Computers & Security*, 90:101691.

Christidis, K. and Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, 4:2292–2303.

Di Martino, B., Mastroianni, M., Campaiola, M., Morelli, G., and Sparaco, E. (2019). Semantic techniques for validation of gdpr compliance of business processes. In *Conference on Complex, Intelligent, and Software Intensive Systems*, pages 847–855. Springer.

Feng, C. Q. and Wang, T. (2019). Does cio risk appetite matter? evidence from information security breach incidents. *International Journal of Accounting Information Systems*, 32:59–75.

Frattolillo, F., Landolfi, F., and Marulli, F. (2009). A novel approach to drm systems. In *2009 International Conference on Computational Science and Engineering*, volume 3, pages 492–497. IEEE.

Karpiriski, M., Senart, A., and Cahill, V. (2006). Sensor networks for smart roads. In *Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06)*, pages 5 pp.-310

Landoll, D. J. and Landoll, D. (2005). *The security risk assessment handbook: A complete guide for performing security risk assessments*. CRC Press.

Maple, C. (2017). Security and privacy in the internet of things. *Journal of Cyber Policy*, 2(2):155–184.

Marulli, F., Pareschi, R., and Baldacci, D. (2016). The internet of speaking things and its applications to cultural heritage. In *IoTBD*, pages 107–117.

Marulli, F. and Vallifuoco, L. (2017). Internet of things for driving human-like interactions: a case study for cultural smart environment. In *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*, pages 1–9.

Mena, D. M., Papapanagiotou, I., and Yang, B. (2018). Internet of things: Survey on security. *Information Security Journal: A Global Perspective*, 27(3):162–182.

Mercaldo, F. and Santone, A. Deep learning for image-based mobile malware detection. *Journal of Computer Virology and Hacking Techniques*, pages 1–15.

Odiete, O., Lomotey, R. K., and Deters, R. (2018). Using blockchain to support data and service management in IoV/IoT. In Peng, S.-L., Wang, S.-J., Balas, V. E., and Zhao, M., editors, *Security with Intelligent Computing and Big-data Services*, pages 344–362, Cham. Springer International Publishing.

Parker, D. B. (1983). *Fighting computer crime*. Scribner New York, NY.

Reyna, A., Martín, C., Chen, J., Soler, E., and Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, 88:173–190.

Sattarova Feruza, Y. and Kim, T. (2007). It security review: Privacy, protection, access control, assurance and system security. *International journal of multimedia and ubiquitous engineering*, 2(2):17–32.

Shameli-Sendi, A., Aghababaei-Barzegar, R., and Cheriet, M. (2016). Taxonomy of information security risk assessment (isra). *Computers & security*, 57:14–30.

Solangi, Z. A., Solangi, Y. A., Chandio, S., bt. S. Abd. Aziz, M., bin Hamzah, M. S., and Shah, A. (2018). The future of data privacy and security concerns in internet of things. In *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, pages 1–4.

Sterbenz, J. P., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M., and Smith, P. (2010). Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*, 54(8):1245–1265.

Team, I. P. (2017). *EU general data protection regulation (GDPR): an implementation and compliance guide*. IT Governance Ltd.

The European Union (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*.

Truong, N. B., Um, T., Zhou, B., and Lee, G. M. (2018). Strengthening the blockchain-based Internet of Value with trust. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–7.

van Geelkerken, F. W. J. and Konings, K. (2017). Using blockchain to strengthen the rights granted through the GDPR. In 7th International youth science forum "Litteris et Artibus", pages 458–461.

Voigt, P. and Von dem Bussche, A. (2017). The eu general data protection regulation (gdpr). *A Practical Guide, 1st Ed., Cham: Springer International Publishing*.

Wang, X., Zha, X., Ni, W., Liu, R. P., Guo, Y. J., Niu, X., and Zheng, K. (2019). Survey on blockchain for internet of things. *Computer Communications*, 136:10–29.

YANG, J., KIM, C., and ONIK, M. M. H. (2019). Aggregated risk modelling of personal data privacy in internet of things. In *2019 21st International Conference on Advanced Communication Technology (ICACT)*, pages 425–430.

Zheng, X., Zhu, Y., and Si, X. (2019). A survey on challenges and progresses in blockchain technologies: A performance and security perspective. *Applied Sciences*, 9(22):4731.